This document is provided as a sample research strategy. Some text has been redacted.

RESEARCH STRATEGY

1. Project Science Areas

2 CB; 6 MCB.

2. Project Description

2.1 Scope

This project will develop a strategy to discover plant natural products (NPs) of potential medicinal value and their biosynthesis through reprogramming the plant innate immunity. This strategy will be rationalized and established through (1) the establishment of a yeast platform that enables functional reconstitution of plant immune complex and high-throughput phenotyping; (2) engineering plant immune receptors so that the chimeric receptors can be activated by known stimulus; and (3) implementation of the chimeric immune receptors into a model plant for plant NP and biosynthesis discovery. The innovativeness of the proposed project is in its utilization of a unique strategy—engineering plant pattern recognition receptors (PRRs) to activate the downstream signaling and metabolic pathways, putatively including a number of secondary metabolic pathways—for the discovery of novel plant NPs and biosynthetic pathways. The expected outcomes of this project include (1) discovery of novel plant NPs that may not be synthesized under normal conditions, the biosynthesis, and the native functions; (2) strategies for engineering plant PRRs that can activate various plant immune and metabolic responses; and (3) insights into how the highly complex plant signaling pathways (immune, perception, growth factor, plant hormone, etc.) are correlated and regulated. The development of this strategy is a significant step towards my long-term career goals: (1) to advance the foundational understanding of phytochemical synthesis in nature, (2) to promote the discovery of novel phytochemicals for agricultural and pharmaceutical applications, and (3) to develop microbial bio-production of plant NPs of high importance and value as an economic approach.

2.2 Background and Introduction

NPs are bioactive compounds isolated from nature that exhibit a wide range of bioactivities. They are of great pharmaceutical significance and play an indispensable role in modern drug development—NPs and their derivatives comprise over half of the current small molecule drugs. Plants, as one major source of NPs, have been used by humans for medicinal applications as early as 2600 BC. To date, phytochemicals still play an important role in the discovery and development of novel small molecule drugs, with more than 10% of the WHO listed essential medicines being of flowering plant origin. With the dramatic development of next-generation sequencing techniques, the genome and transcriptome data of a large number of organisms are now available. Analysis of the genome data implies that nature's synthetic potential has been largely underestimated, and the majority of NPs have been unexplored. On the other hand, in recent decades, there has been a decline in NP-based drug development. This is in part due to the challenges associated with decreased frequency of identifying novel NP structures with traditional methods. Many approaches have been developed in the recent twenty years to more efficiently discover novel NPs from microbial organisms. Two major strategies have been developed to discover the previously uncharacterized NPs: (1) prediction of biosynthetic gene cluster that may synthesize novel structures from genome sequence information, followed by functional verification in either native or heterologous host, and (2) introduction of various stimuli to the native producing host to induce the expression of the cryptic biosynthetic pathways.

However, none of these strategies are applicable to predicting or activating biosynthetic pathways that are cryptic under normal conditions in plants. Unlike their microbial counterparts, genes encoding the biosynthetic enzymes are believed to be primarily distributed throughout the plant genome. Thus, while enzymes responsible for the synthesis of microbial NPs are normally discovered and characterized in batches, those associated with plant NPs are typically discovered individually. This has contributed to the disparity in understanding of microbial versus plant NP biosynthesis. And the pool of plant biosynthetic enzymes is not large enough for a clear sequence-function correlation to be established and serve as a reference during bioinformatics analysis of the genomic and transcriptomic data. On the other hand, introducing stimuli is a relatively difficult-to-achieve strategy in plant because plants are higher multicellular organisms. Also, due to the unpredictability of this method, the discovery of novel molecules using this strategy will be highly inefficient. Thus, an alternative strategy to activate plant secondary metabolic pathways is necessary to more efficiently discover novel plant NPs that cannot be detected or isolated under normal cultivation conditions.

It has been widely believed that plant natural products play an important role in plant defense against pathogens and herbivores. Pattern recognition receptors (PRRs) are surface-localized, ligand-binding, receptor-like kinases or proteins (RLKs or RLPs) that recognize and respond to molecular signals in the environment. PRRs are central to immune-signaling as well as other pathways such as hormone-signaling and stress-signaling. Genomic

analysis indicates that the model plants *Arabidopsis thaliana* and rice each encode more than 600 PRRs, and the functions of >95% of these putative PRRs remain unknown. Each PRR putatively will responds to certain stimuli (detrimental or beneficial, exogenous or endogenous) for plant to generate development and environmental responses (either abiotic or biotic stress). If we can activate the downstream pathways of each PRR, then plant NPs that are produced as defense or signaling molecules under various situations will be activated and discovered.

Therefore, we propose to develop a strategy to engineer and redirect plant perception to activate and discover novel plant NPs. This strategy will not only advance the discovery of novel plant NPs and the biosynthesis, but also shed light on the fundamental understanding of the functions of plant PRRs and NPs. Our central hypotheses are (1) plant perception complex/immune signaling pathway can be functionally reconstituted in the backer's yeast *Saccharomyces cerevisiae*; (2) plant PRRs can be engineered to recognize altered stimulus; and (3) the signaling and metabolic pathways downstream of target plant PRRs can be activated through introducing the engineered PRR and the corresponding stimulus. With expertise in synthetic biology and NP biosynthesis, we will verify the hypotheses and demonstrate this strategy through executing the following three specific aims, using *A. thaliana* and tobacco as the testing plant (Fig. 1):

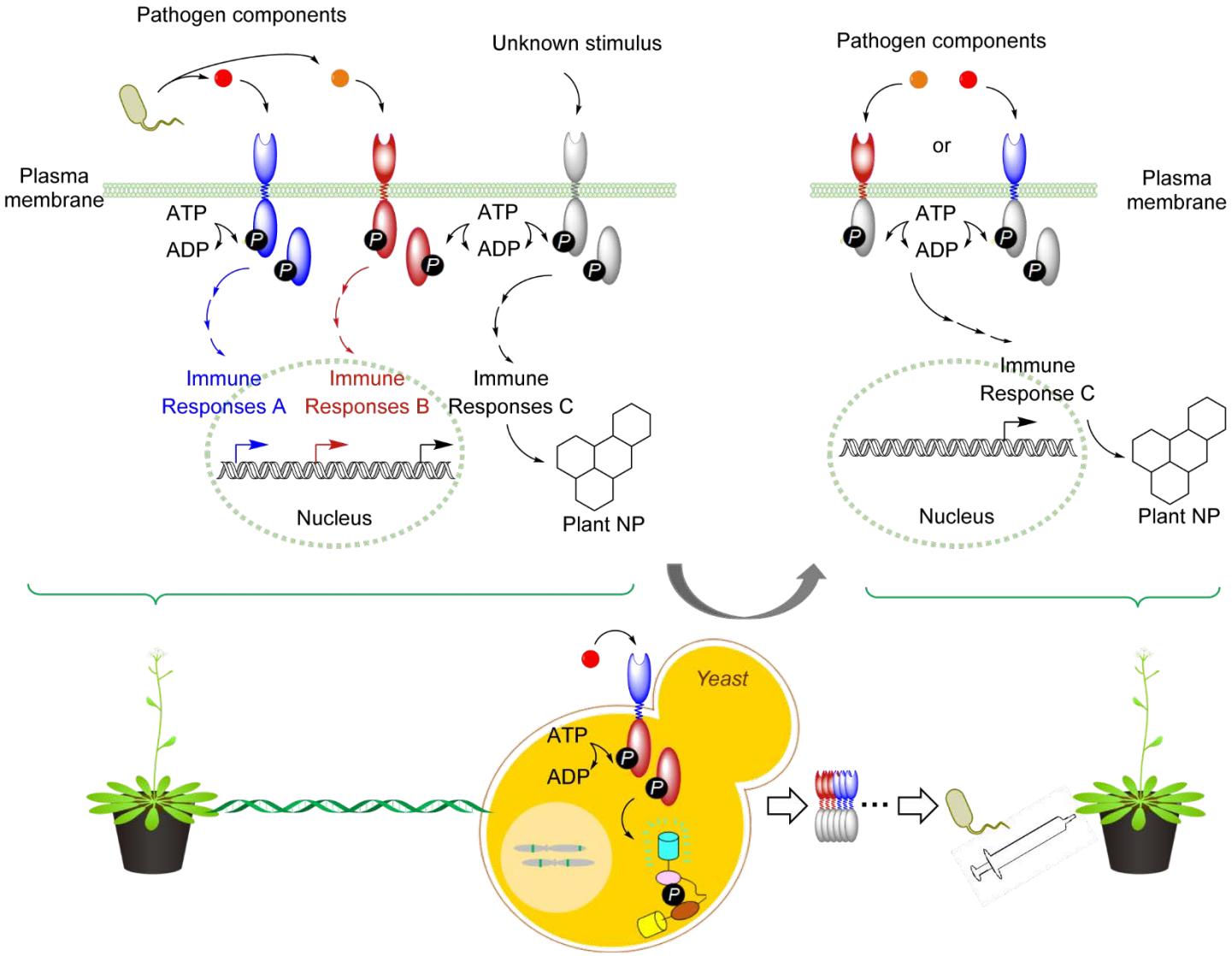

Fig. 1 Illustration of the proposed project on plant NP discovery through redirecting plant immune signaling.

Aim 1: *Establish a yeast platform that enables functional reconstitution of plant immune complex and high-throughput phenotyping.* A set of kinase sensors that can indicate the phosphorylation status of target kinase will be developed in yeast; and the best-studied plant immune complex will be functionally reconstituted in yeast. Incorporation of the kinase sensor with the reconstituted plant immune complex will enable high-throughput phenotyping of the PRR function.

Aim 2: *Engineer plant immune receptors so that they can be activated by known stimulus.* Domain swapping will be employed to engineer PRRs to recognize the altered elicitor. The boundary for the chimera domain swapping to make the chimeric PRR will be examined and ruled out for different types of PRRs tested in the project. A library of chimeric PRRs will be generated from this specific aim, with a focus on leucine rich repeats (LRR) receptor kinase and proteins.

Aim 3: *Implement the chimeric immune receptors into model plant for plant NP and biosynthesis discovery.* The chimeric PRRs from ***Aim 2*** will be introduced into *A. thaliana* through agroinfiltration with or without the inactivation of the related native PRRs. Metabolomics and transcriptomics will be applied to detect the production of unknown natural products and the enzymes putatively involved in the biosynthesis. We will also try to reconstitute the activated pathway predicted from the transcriptome data in yeast to elevate the production level and characterize the identity of related metabolites.

2.3 Rationale

This project is proposed based on the following facts:

Plant secondary metabolites are believed to play an indispensable role in plant innate immunity and defense framework. Collectively, plants can produce more than 100,000 NPs of diverse structures and bioactivities [1]. Phytochemicals involved in plant defense are divided into two groups, phytoanticipins and phytoalexins. Phytoanticipins are compounds constitutively produced and stored in plants. Phytoanticipins are often present in plants in their inactive form, and enzymes converting these molecules to the bioactive forms are only expressed in response to the corresponding stimuli. On the other hand, the biosynthesis of phytoalexins are not constitutive and only induced upon the recognition of pathogen or damage, i.e. the biosynthesis of the inducible phytoalexins are closely correlated to the recognition of pathogens or damage through PRRs [2]. Thus, if the many of the PRRs of unknown functions can be engineered to be activated by known elicitors, the biosynthesis of phytoalexins that are linked to these unknown PRRs will be induced, thereby enabling the discovery of novel phytochemicals. Additionally, while phytoanticipins are comparatively better investigated, the proposed strategy will promote the discovery and fundamental understanding of phytoalexins, biosynthesis, and their native functions.

Domain swapping strategy has been employed to uncover the native function of unknown plant PRRs. Although different PRRs are composed of different domain architectures, they mostly contain an ectodomain, such as an LRR domain, lectin domain, lysine motif, and more. Chimeric receptors constructed through domain swapping has been a useful approach to elucidate the function of unknown receptors in animals. This strategy has been successfully demonstrated recently in plants, but with far fewer examples. The first chimeric receptor was constructed through fusing the ectodomain of [redacted] and kinase domain of [redacted] which perceives [redacted] inducer brassinolide and initiates defense responses downstream [redacted]. More recently, another example of fully functional chimeras was constructed between [redacted] Both [redacted] ectodomain: [redacted] kinase domain chimera and EFR ectodomain:WAK1 kinase domain chimera function as expected. [redacted] are all LRR receptor kinases, while WAK1 is a pectin receptor kinase. Thus, fully functional chimeric receptors can be obtained not only from two structurally similar RLKs, but also from two distinct RLKs. LRRs are protein structural motifs that normally participate in protein-protein interactions and are heavily employed in plant innate immunity through binding to the pathogen or damage associated molecular patterns to perceive pathogen or damage. In *A. thaliana*, 239 of the PRRs contain leucine-rich repeats (LRRs) in the extracellular domain. In addition, the crystal structures of plant LRR PRRs and their interaction have been achieved and analyzed, which provides more guidance on construction of the chimeric PRRs. Thus, this project will focus on engineering LRR receptor kinases or proteins.

Yeast is a desired platform to investigate and engineer plant receptors. *S. cerevisiae* has been an essential model organism for studying eukaryotic cell biology, was it has well-established genetic modification tools. Previous systematic comparison of protein kinases between *Arabidopsis* and *S. cerevisiae* indicates that *S. cerevisiae* does not harbor homologous RLK or RLP-associated signaling pathways to *Arabidopsis*. When we investigate the plant immune system through reconstructing the plant signaling machineries orthogonally in yeast, it will present a “clear-cut” view on the regulation and function of plant RLK/RLPs. This strategy also enables the development of high-throughput phenotyping strategy because (1) yeast is unicellular and (2) there are a number of tools and sensors available in yeast.

Yeast has been widely used as a platform for plant biosynthetic pathway reconstitution and identification. Yeast systems are often used in the heterologous production of plant metabolites because they are evolutionarily

closer to the plant. *S. cerevisiae* is the most common choice when using yeast to synthesize plant molecules, largely because *S. cerevisiae* is one of the best-studied and most genetically accessible microorganisms. Many highly complex plant biosynthetic pathways have been successfully reconstructed and identified in baker's yeast, such as medicinal opioid alkaloids, precursor to the antimalarial artemisinin, and precursors to vinblastine, and many more. In our preliminary studies, we have demonstrated that (1) the plant biosynthetic pathway can be comprehensively elucidated and reconstituted at the same time using the heterologous yeast strain [5] and (2) *S. cerevisiae* can be used to express highly complex plant biosynthetic pathways efficiently[6]. Based on these results, we believe that *S. cerevisiae is a powerful heterologous host for studying and engineering plant NP biosynthesis.*

3. Research Plan

Aim 1: Establish a yeast platform that enables functional reconstitution of plant immune complex and high throughput phenotyping.

Plants activate immunity by recognizing pathogen-generated elicitors or damage associated molecular patterns using PRRs, which are surface-localized ligand-binding RLKs or RLPs. In 1994, CF-9 was successfully cloned from tomato as the first identified RLP. [redacted]

Functional reconstruction of a well-elucidated model plant RLK complex in *S. cerevisiae*. To demonstrate that yeast provides the required microenvironment for the functional heterologous expression of plant RLKs, a well-characterized plant RLK complex will be reconstituted in yeast. We choose the [redacted] as the model pathway to reconstitute in yeast. The coding regions of [redacted] have each been cloned into the low-copy yeast expression cassettes. Each enzyme will be tagged with N- and C-terminus HA tag to check the heterologous expression through Western blotting, and EGFP tag to verify the correct localization in yeast through microscopy. [redacted] are expected to localize on the cell membrane, while [redacted] is anticipated to be detected mainly in the cytosol. Then, [redacted] will be co-expressed in yeast, and the auto- and transphosphorylation with or without the presence of [redacted] will be examined using Mass Spectrum analysis and SDS PAGE analysis with Phos-tag™.

Establishment of a kinase sensor in yeast. Recently, a FRET-based kinase biosensor has been established as a real-time monitoring tool to track the activity of [redacted] in mammalian cells. We have reconstituted the function of FRET-based [redacted] in yeast [redacted]. We have also converted the FRET-based FAK sensor into a yeast-two-hybrid (Y2H) based sensor in yeast using an auxotrophic marker as the reporter gene. [redacted] have been identified as the transphosphorylation substrate by the [redacted] complex. The peptide substrate of [redacted] will then be replaced with the peptide sequence from [redacted] in the middle [redacted] and each introduced into the engineered yeast strain expressing [redacted] to monitor the transphosphorylation activity of the complex. The phosphorylated peptide binding domain of the kinase sensor will be engineered for optimal performance (stronger affinity to the phosphorylated peptide substrate) in yeast. Additionally, we will also employ yeast-two-hybrid system to establish a second type of kinase sensor, with the putative substrate expressed on the C-terminus of the activation domain and the phosphorylated tyrosine or serine-binding domain expressed on the C-terminus of binding domain of GAL4. Once AD-peptide substrate is phosphorylated, it will interact with the phosphorylated tyrosine or serine binding domain, and thus initiate the expression of reporter gene (luciferase or auxotrophic selection marker).

Functional reconstitution of additional PRR systems in yeast. Upon the verification of the [redacted] being functionally reconstituted in yeast with the kinase sensor, additional model PRR systems will be reconstituted. Although previous studies have demonstrated the possibility and potential to construct fully functional chimeric PRRs from two structurally distinct PRRs (LRR and lectin receptor kinases), there are many types of PRRs, and functional chimeric PRRs might not be able to be functionally constructed from certain types of PRR with the well-studied LRR receptor kinases. We aim to cover as many types of PRR systems as possible for the subsequent engineering in ***Aim 2***, with a focus on best-studied LRR RLKs. The model PRR systems to be constructed in yeast is listed in [redacted] with the target PRR underlined.

Expected Outcome: Functional reconstruction of the multiple PRR complex with the corresponding kinase sensor in yeast. The yeast platform not only enables subsequent establishment of functional chimeras, but also allows

high-throughput phenotyping to identify ligands or missing components to functionally reconstruct the target PRR complex in yeast.

Aim 2: Engineer unknown plant immune receptors so that they can be activated by known elicitor.

PRRs are composed of two domains: ectodomain domain responsible for the elicitor recognition, and kinase or signaling domain responsible for the initiation of downstream signal transduction cascade. Several chimeric receptors have been previously constructed, such as the [redacted] chimeras [3, 4]. We will swap the ectodomain of the unknown PRRs with the corresponding domain of PRRs established in ***Aim 1***. The recognition spectrum of the chimeric PRRs will be altered (can be activated by the elicitors used in ***Aim 1***) and at the same time will induce the signaling/metabolic pathways downstream of the unknown PRRs [8]. The detailed domain swapping strategy will be examined through constructing an array of chimeric PRRs between the well-characterized PRRs reconstituted in yeast in ***Aim 1***. Because previous examples on chimeric PRRs are all related to LRR RLKs, LRR RLKs are well characterized and play an important role in plant innate immunity. Therefore, constructing chimeric PRRs between unknown LRRs and well-characterized LRR RLKs or RLPs will be the focus in ***Aim 2***.

Reconstruction of the [redacted] chimeras in the yeast-based system. To demonstrate yeast as a feasible platform for the functional verification of the chimeric PRRs, we will first reconstruct the previously reported [redacted] chimeras. The chimeric PRR composed of [redacted] ectodomian and [redacted] kinase domain will be examined on the yeast strain harboring the [redacted] system, with the kinase sensor monitoring the activity of [redacted]. [redacted] will be either replaced with the chimera or retained to examine the effect of the wildtype [redacted] on the engineered [redacted] chimera in yeast. Similarly, the chimera composed of [redacted] ectodomain and [redacted] domain will be constructed as previously reported, reconstituted, and verified in yeast [redacted].

Establishment of the domain swapping strategy between the characterized and unknown LRR RLKs in yeast. Previous study on [redacted] chimera indicates that only the chimera constructed through domain swapping between intracellular juxtamembrane (JM) domain and kinase domain is functional [3]. On the other hand, study on [redacted] implies that domain swapping between extracellular JM domain and transmembrane (TM) domain leads to higher activity [4]. Here, on the basis of the yeast platform, we will test different boundaries for the domain swapping between any two of [redacted] to establish a general rule for choosing the boundary for domain swapping to obtain a fully functional chimeric PRR. The flexibility of the yeast platform will allow us to efficiently test the effect of all the possible boundaries on activities of six chimeras [redacted]. Upon the establishment of the domain swapping strategy, three sets of chimeras will be constructed between the ectodomain of the three characterized LRR RLKs and the kinase domain of the other unknown LRR RLKs.

Construction of the chimeric receptors between the characterized LRR RLKs and unknown LRR RLPs in yeast. Different from LRR RLK, there is a group of LRR RLPs that do not contain kinase domains, but also play an indispensable role in plant immunity, e.g., the comparatively well-characterized [redacted] that are responsible for the recognition of fungal pathogens in tomato. Due to the relatively simpler activation mechanism of [redacted] (i.e. addition of [redacted] respectively), we will construct chimeras between the ectodomain of these three RLKs and the remaining part (extracellular JM, TM domain and cytoplasmic region) of [redacted]. Similarly, different boundaries will be chosen, and the function of the chimeras will be verified, using the kinase sensor targeting [redacted] respectively. Due to the structural similarity between LRR RLK and LRR RLP (kinase domain in RLK while cytoplasmic region in RLP), we expect the boundary for domain swapping to be similar to the LRR RLK chimera construction.

Construction of other chimeric receptors between the three LRR RLKs and other types of PRRs. Similarly, chimeras between the ectodomain of LRR RLKs [redacted] and other type of PRRs will be constructed following the rule of boundaries in chimeric LRR RLK or RLP construction. The function will be verified by the yeast strain harboring the corresponding immune complex and the kinase sensor if applicable. If the chimeras are not as active, different boundaries will be tested and examined. If the chimeras keep failing to exhibit activity, it is likely that the ectodomain of LRR RLKs is not compatible with the specific group of PRRs. A different ectodomain will be employed[redacted]. Upon the verification of the domain swapping strategy, one or multiple set(s) of chimeras will be made from the known ectodomain and the kinase domain of unknown PRRs.

Expected Outcome: Rule out the domain swapping strategy on the yeast platform for different types of PRRs (with a focus on LRR RLKs or RLPs) and generate an array of plant chimeric PRRs that can be activated by known elicitors.

Aim 3: Implement the engineered immune receptors into the model plant for plant NP and biosynthesis discovery.

We hypothesize that the engineered PRRs constructed in ***Aim 2*** will activate the signaling and metabolic pathways downstream of the kinase domain or cytoplasmic region in response to the presence of the stimulus to the ectodomain of the chimeras. Thus, we will implement these chimeric PRRs in the model plant *A. thaliana* and tobacco, to see if we can discover plant natural products of novel structures through this proposed strategy.

Verification of the compatibility of the yeast platform through testing the chimeras built from characterized PRRs in the plant. To confirm that the yeast platform is compatible with the plant system (i.e., the chimeric PRRs that show activity in the yeast platform are also active in the plant), all the chimeras that are made from the well-characterized PRRs [redacted] will be introduced into *A. thaliana* through agroinfiltration. Since [redacted] the plant hormone brassinolide, which will lead to a number of development responses, we will not use the chimera containing the ectodomain of [redacted]. Elicitors to the corresponding ectodomain of the chimeras will be applied on the agroinfiltrated patches, subsequent to the verification of the activation status of the pathways downstream of the kinase domain. For example, if the chimera contains the [redacted] kinase domain, we will verify the activity of this chimera on the downstream pathways by measuring the ethylene abundance, which should be levitated in response to the corresponding elicitors. In addition, we will use these chimeric PRRs to examine if it is necessary to inactivate or downregulate the expression of the native PRRs that contribute to the ectodomain. Because the ectodomains we are using in the chimeras are mostly the well-studied ubiquitous PRRs [redacted], it is very likely that in response to the addition of the elicitors, the native PRRs will affect the function of the chimeric PRRs. RNA-Seq will be applied to estimate the effect of the elicitor and the function of the chimeric PRRs.

Heterologous expression of the chimeric PRRs in *Arabidopsis*. Once the functional chimeric PRRs constructed in yeast are confirmed to be functional in plants, the different set of chimeric PRRs will be introduced into *A. thaliana* and tobacco through agroinfiltration. The chimeric LRR RLKs will be examined first, because this group is comparatively better studied, and there are two versions of chimeras for each unknown PRR (with ectodomain of [redacted]). According to the aforementioned investigation on the effect of the native [redacted] to the chimeras, we will construct an [redacted] double knockout *Arabidopsis* or tobacco using CRISPR technology or downregulate the expression level of [redacted] in *Arabidopsis* or tobacco through vitrus induced gene-silencing technique. Agrobacterium holding the corresponding chimeric PRR or the wild type unknown LRR RLK (negative control) expression cassette will be introduced into the [redacted] or the wild type *A. thaliana* and tobacco. The corresponding elicitor [redacted] or just the blank solution with no elicitors will be applied on the agroinfiltrated patches in triplicates, incubated for one week, with phenotype recorded, then half lyophilized for metabolite extraction with water followed by ethanol, and the other half liquid nitrogen treated for RNA extraction. The extracted RNA will be used for the preparation of the RNA-Seq libraries, and the extracted metabolites will be analyzed and characterized by reverse phase LC-MS/MS on a Waters Synapt G2-Si Q-TOF (ACQUITY UPLC CSH Phenyl-Hexyl Column, 130Å, 1.7 µm, 2.1 mm×100 mm) using both positive and negative ionizations at the metabolomics core facility. Metabolites will be separated on a linear gradient of 1% acetonitrile (v/v in water, 0.1% formic acid) to 90% acetonitrile (v/v in water, 0.1% formic acid) over 30 min with a flow rate of 0.25 ml min^{-1}. [redacted] the specific receptors inducing the synthesis of this defense molecule have not been clearly elucidated. [redacted] and at the same time target to clearly illustrate the set of receptors that are linked with the induction of the camalexin biosynthesis. If new compounds are detected from metabolomics analysis, the corresponding receptor will be transiently expressed in plant on a larger scale, so that we will have enough crude lysate for compound extraction, purification, and NMR analysis for structural elucidation. RNA-Seq data of the chimeric PRR expressing plant and the negative control will be compared and analyzed to propose candidate enzymes that are putatively involved in the biosynthesis, which will be reconstituted and verified in yeast (parallel with other ongoing projects in my laboratory). Similarly, other groups of chimeric PRRs will be introduced into plants followed by transcriptomics and metabolomics. We will also compare the metabolomics and transcriptomics data between *Arabidopsis* and tobacco.

Expected Outcome: Identification of the set of receptors that initiate the biosynthesis of camalexin in *Arabidopsis* and tobacco. Uncover novel phytochemicals synthesized from *A. thaliana* and tobacco, elucidate the biosynthesis, and propose their biological function. At the same time, propose putative pathways upregulated in *A. thaliana* and tobacco upon elicitation, and identify the function and corresponding products through synthetic biology strategy in yeast.

4. Innovativeness

This proposed project is innovative from three aspects.

First, although there have been quite a number of various strategies developed to discover NPs, few are about NP discovery in plants. The current strategies in NP discovery mainly employ two mechanisms: genomics-based prediction of cryptic gene clusters that may synthesize novel structures, or introducing various stimuli to activate the gene clusters that are cryptic under normal lab conditions (e.g., culture the producer strain with other competing strains). However, both mechanisms are less applicable in the case of plant NP discovery. Our proposed strategy circumvents these challenges and seeks a distinct strategy for activating these cryptic pathways under normal conditions. Domain engineering of immune receptors is an established yet relatively new method to elucidate the function of unknown plant PRRs. Thus, the results from this project will also shed light on strategies for engineering plant immune receptors, which can serve as a general strategy to uncover the functions of many unknown PRRs in plants.

Second, while previously constructed chimeric PRR was examined in plants, this project proposes to construct and verify the functional chimeric PRR in yeast. In the previous example constructing plant chimeras, the functions of the chimeric PRRs were verified with indirect phenotype [4]; by contrast, in the proposed project, the function of the chimeric PRRs will be verified directly through the activity of the kinase domain. In addition, the yeast platform will enable more efficient construction of various chimeras due to the more efficient genetic engineering tools and high-throughput phenotyping capability of *S. cerevisiae*.

Third, this strategy will provide a unique source to better understand the functions of the relating plant NPs in their native producing host, which are normally difficult to uncover. Most plant NPs function as signaling or defense molecules, but the exact functions are often unknown. Each plant receptor is responsible for recognizing certain stimulus and is in charge of turning on and off a number of downstream signaling and metabolic pathways. Thus, transcriptome analysis of the engineered plant not only can facilitate the identification of enzymes involved in the biosynthesis of interesting compounds, but will also hint at signaling or metabolic pathways (mainly the already characterized) that are correlated with the produced compounds of interest, and thus shed light on the native functions of the target plant NP. Additionally, this strategy will generate engineered plants that can efficiently produce the desired compounds, the development of which from scratch is usually highly challenging and time-consuming.

On the other hand, this proposed project is also high-risk.

First, the model plant *A. thaliana* and tobacco may produce a very limited number of phytoalexins—camalexin is believed to be the major phytoalexin in *Arabidopsis*. Thus, we will also test our strategy in tobacco, which is known to produce more NPs including phytoalexins. However, if we obtain very similar metabolic profiles from different chimeric PRRs in these two model plants, we will switch to other model plants, preferably tomato and rice, which are known to produce a variety of alkaloids and terpenoids.

Second, only chimeras built from LRR and lectin receptor kinases have been successfully demonstrated, but there are many other types of plant PRRs not yet examined or demonstrated. The best characterized LRR, lectin, lysM receptor kinases will be the focus of this project—they play an important role among plant PRRs, e.g. in *A. thaliana*, among the 611 PRRs encoded by the genome, 239 are LRR receptor kinases/proteins, 46 are lectin receptor kinases, and 4 are LysM receptor kinase [8]. The remaining PRRs, from which functional chimeras cannot be constructed, will be reconstituted in the yeast platform, and tested for the possibility of screening for peptide or small molecule ligands that can activate the kinase activity of the target PRR, from peptide or small molecule libraries. A generic kinase sensor instead of a substrate specific kinase sensor will be developed and used to screen for the ligands. If a ligand can be selected from the ligand library, it will be applied on the native plant host followed by omics analyses.

Third, plant immune response pathways have been recognized to be highly intertwining with other signaling pathways such as plant hormone signaling pathways. Thus, the function of the chimeric PRR may be influenced by the native PRRs. If that happens, we will test the chimeric PRR in plants with the two native PRRs inactivated or downregulated and enhance the expression level of the chimeric PRRs. Moreover, the immune complex we choose may not be the intact module (minimal components to remain functional)—there are other enzymatic components missing from the complex [redacted]—and that inhibits the complex from being fully functional in yeast. In addition to introducing more putative components to reconstitute the activity (e.g. for [redacted] we will try the phosphatase [redacted] that may dock with [redacted] as an inactive form), the missing component may also be screened from the plant cDNA using the corresponding kinase sensor in the yeast-based system.

5. Investigator Qualifications

The goal of the proposed research is to discover novel plant NPs through engineering and reprogramming the plant immune complex. The development of this strategy is a significant step towards my long-term career goals: (1) to advance the foundational understanding of phytochemical synthesis in nature, (2) to promote the discovery of novel phytochemicals for agricultural and pharmaceutical applications, and (3) to develop microbial bio-production of plant compounds of high importance and value as an economic approach. I have extensive background, expertise, and leadership skills to serve as PI and direct the proposed activities to success.

I've been working on the discovery, elucidation, and engineering of plant NPs and biosynthesis for more than ten years. My training spans chemical engineering, chemistry and biology, allowing me to effectively integrate core concepts of engineering (such as design, analysis, control) with the substrate of biology. My graduate training focused on using engineering strategies (mainly protein engineering and metabolic engineering) to decipher the sequence-function correlation of the key multi-domain mega-synthase (non-reducing polyketide synthase composed of up to seven domains) in the synthesis of aromatic polyketides in filamentous fungi. My postdoctoral training focused on harnessing synthetic biology in yeast to elucidate and engineer complex plant pathways (specifically, benzylisoquinoline alkaloid biosynthetic pathway). Currently, the ongoing projects in my laboratory include: (1) deep mapping of plant apo-carotenoid and phytosterol biosynthesis in yeast; (2) developing synthetic biology tools in yeast to enhance the bioproduction of plant NPs; (3) reconstructing, engineering, and discovery of plant immune complex in yeast. These projects enable the establishment of collaboration of my laboratory with laboratories of plant pathology, plant biology, bioinformatics, and biochemistry, and also set the basis for conducting the proposed project (Fig. 4).

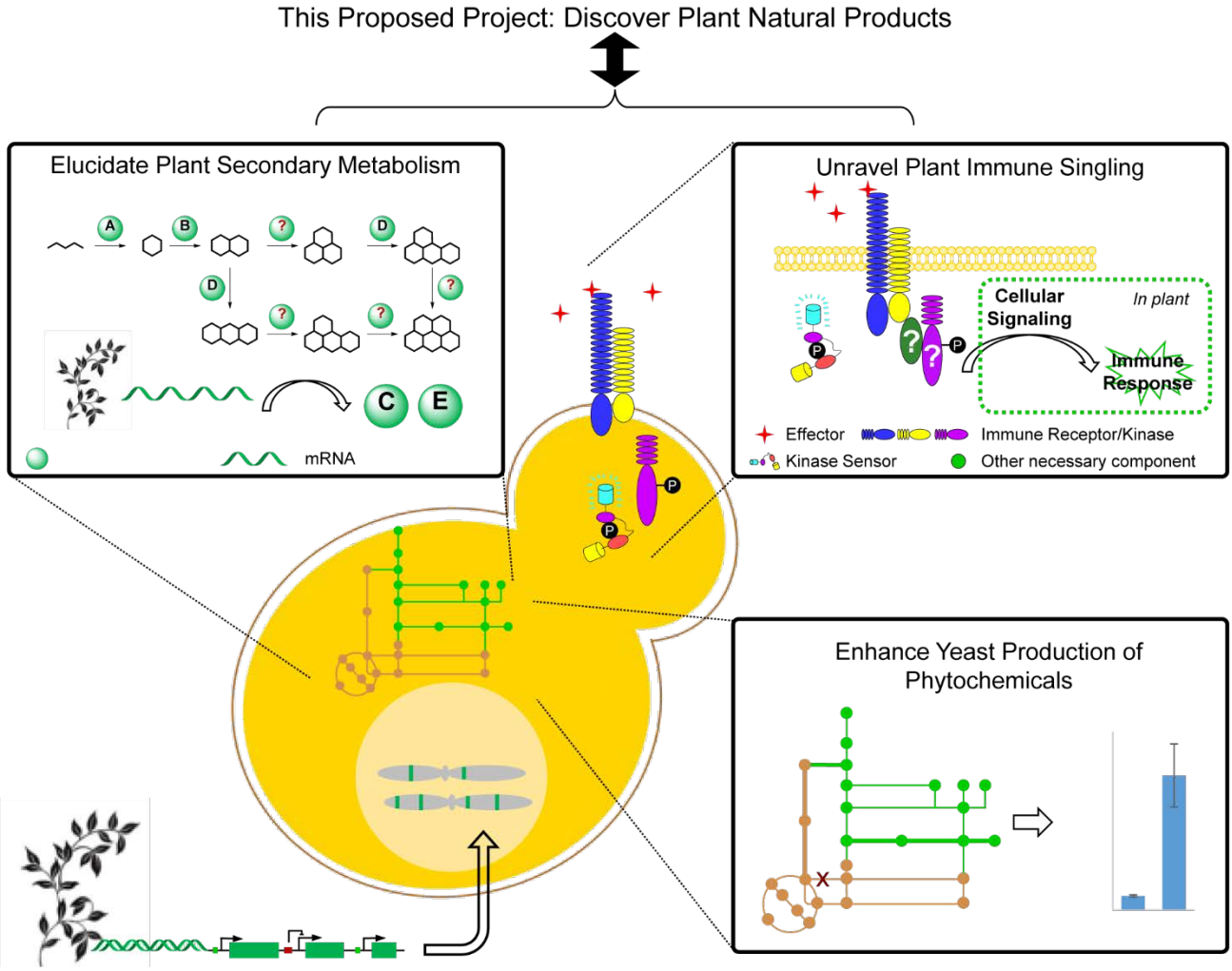

Fig. 4 Illustration of the current ongoing projects in my laboratory. This proposed project is well incorporated with these projects and is a significant step towards my long-term career goals.

In addition, I also have a demonstrated record of successful, high-impact, and productive research projects that includes several previously unprecedented engineering efforts, such as (1) the first bioproduction of bacterial aromatic polyketides (an important group of NPs including the broad spectrum antibiotics tetracyclines and the

anticancer anthracyclines) in the workhorse *E. coli*, (2) one of the first functional reconstitutions of fungal biosynthetic pathways in yeast, (3) pioneering in establishing sequence-function correlations for fungal NP biosynthetic enzyme to predict the function of uncharacterized enzymes (specifically, non-reducing polyketide synthases), (4) the first demonstration of a complex plant biosynthetic gene cluster elucidation using synthetic biology in yeast, and (5) successful and efficient reconstitution of the most complicated plant biosynthetic pathway in yeast. Although engineering plant immune signaling is a new topic to me, my previous training has laid a solid foundation for the proposed engineering efforts in reconstructing, elucidating, and engineering complex plant pathways (specifically, plant immune signaling pathways) in yeast. Therefore, my technical background and expertise make me well-suited to lead the proposed activities.

6. Suitability for the New Innovator Award Program

This proposed project, as previously discussed, is innovative yet at the same time highly challenging—few chimeric receptor examples have been successfully demonstrated in plants, compared with in animals. In addition, the development of this strategy is a significant step towards my long-term career goals, and this proposed project is complementary to my ongoing research projects which include characterization and engineering of plant NP biosynthesis that also sets the basis for the proposed project. On the other hand, the proposed work is a new extension from my current research projects in that it involves functional reconstitution and engineering of the plant immune complex in both yeast and plants. If this strategy can be successfully demonstrated, it will not only accelerate the discovery of plant NPs but will also help us understand the native function of these valuable compounds in plants, which is important knowledge to better predict and discover novel plant NPs. However, the proposed approach (reprograming the plant perception complex) has never been applied on natural product discovery in plants before, and the expected outcome is less predictive than many other proposals. Thus, it is not feasible to apply for a regular grant (e.g. NIH R01, NSF PGRP) to support me to develop this innovative but high-risk strategy.

7. Expected Outcomes and Timeline

This proposed research aims to discover novel plant NPs through engineering and reprogramming plant perception complex. The expected outcomes of this project include (1) discovery of novel plant NPs that may not be synthesized under normal conditions and their native functions; (2) development of strategies for engineering plant pattern recognition receptors that can activate various plant immune and metabolic responses; and (3) enhanced understanding on how the highly complex plant signaling pathways (immune, perception, growth factor, plant hormone, etc.) are correlated and regulated. The timeline of the project is described below.

	Specific Aim	Y1	Y2	Y3	Y4	Y5
1.	Establish a yeast platform for functional reconstitution of plant immune complex and high throughput phenotyping.					
	Functional reconstruction of a model plant RLK complex.	■				
	Establishment of a kinase sensor in yeast.	■	■			
	Functional reconstitution of more model PRR system in yeast.		■	■	■	
2.	Engineer unknown plant immune receptors so that they can be activated by known elicitor.					
	Verification of the yeast platform on chimera function.		■	■		
	Domain swapping strategy between LRR RLKs in yeast.		■	■	■	
	Construction of the chimeric LRR RLPs in yeast.			■	■	■
	Construction of other chimeric receptors.		■	■	■	■
3.	Implement the engineered immune receptors into the model plant for plant NP and biosynthesis discovery.					
	Testing the chimeras built from characterized PRRs in plant.		■	■	■	
	Heterologous expression of the chimeric PRRs in *A. thaliana* and tobacco.			■	■	■

8. Statement of Research Effort Commitment

If awarded, I will commit a minimum of 3 person months (25%) of my research effort to the project supported by this New Innovator Award.

References

1. Dixon RA. Natural products and plant disease resistance. Nature. 2001;411(6839):843-7. Epub 2001/07/19. doi: 10.1038/35081178. PubMed PMID: 11459067.
2. Ahuja I, Kissen R, Bones AM. Phytoalexins in defense against pathogens. Trends Plant Sci. 2012;17(2):73-90. Epub 2012/01/03. doi: 10.1016/j.tplants.2011.11.002. PubMed PMID: 22209038.
3. He Z, Wang ZY, Li J, Zhu Q, Lamb C, Ronald P, Chory J. Perception of brassinosteroids by the extracellular domain of the receptor kinase BRI1. Science. 2000;288(5475):2360-3. Epub 2000/07/06. PubMed PMID: 10875920.
4. Brutus A, Sicilia F, Macone A, Cervone F, De Lorenzo G. A domain swap approach reveals a role of the plant wall-associated kinase 1 (WAK1) as a receptor of oligogalacturonides. Proc Natl Acad Sci U S A. 2010;107(20):9452-7. Epub 2010/05/05. doi: 10.1073/pnas.1000675107. PubMed PMID: 20439716; PMCID: PMC2889104.
5. Li Y, Smolke CD. Engineering biosynthesis of the anticancer alkaloid noscapine in yeast. Nat Commun. 2016;7:12137. doi: 10.1038/ncomms12137. PubMed PMID: 27378283; PMCID: PMC4935968.
6. Li Y, Li S, Thodey K, Trenchard I, Cravens A, Smolke CD. Complete biosynthesis of noscapine and halogenated alkaloids in yeast. Proc Natl Acad Sci U S A. 2018;115(17):E3922-E31. Epub 2018/04/04. doi: 10.1073/pnas.1721469115. PubMed PMID: 29610307; PMCID: PMC5924921.
7. Lu D, Wu S, Gao X, Zhang Y, Shan L, He P. A receptor-like cytoplasmic kinase, BIK1, associates with a flagellin receptor complex to initiate plant innate immunity. Proc Natl Acad Sci U S A. 2010;107(1):496-501. Epub 2009/12/19. doi: 10.1073/pnas.0909705107. PubMed PMID: 20018686; PMCID: PMC2806711.
8. Shiu SH, Bleecker AB. Plant receptor-like kinase gene family: diversity, function, and signaling. Sci STKE. 2001;2001(113):re22. Epub 2001/12/26. doi: 10.1126/stke.2001.113.re22. PubMed PMID: 11752632.